



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification⁶ : C12N 15/82, C07K 14/415, C12N 15/29</p>	<p>A1</p>	<p>(11) International Publication Number: WO 99/47688 (43) International Publication Date: 23 September 1999 (23.09.99)</p>
<p>(21) International Application Number: PCT/US99/05183 (22) International Filing Date: 10 March 1999 (10.03.99) (30) Priority Data: 60/078,144 16 March 1998 (16.03.98) US (71) Applicant (for all designated States except US): E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): CAHOON, Rebecca, E. [US/US]; 2331 West 18th Street, Wilmington, DE 19806 (US). KLEIN, Theodore, M. [US/US]; 2229 Rosewood Drive, Wilmington, DE 19810 (US). (74) Agent: MAJARIAN, William, R.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).</p>		<p>(81) Designated States: AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>
<p>(54) Title: INHIBITORS OF APOPTOSIS PROTEINS IN PLANTS</p> <p>(57) Abstract</p> <p>This invention relates to an isolated nucleic acid fragment encoding an apoptosis inhibitory protein. The invention also relates to the construction of a chimeric gene encoding all or a portion of the apoptosis inhibitory protein, in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the apoptosis inhibitory protein in a transformed host cell.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TITLE

INHIBITORS OF APOPTOSIS PROTEINS IN PLANTS

This application claims the benefit of U.S. Provisional Application No. 60/078,144, filed March 16, 1998.

5

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding proteins that inhibit apoptosis in plants and seeds.

BACKGROUND OF THE INVENTION

10

Multicellular organisms have evolved elaborate signal transduction pathways for maintaining homeostasis through the control of cell proliferation and death. Programmed cell death or apoptosis is a process in which unwanted cells are eliminated during growth and development (Steller, H. (1998) *PNAS* 95:5421-5422). Inhibitors of apoptosis proteins (IAPs) constitute a family of highly conserved death-suppressing proteins that were first identified in baculoviruses (Liston, P. et al. (1997) *Apoptosis* 2(5):423-441). Two homologues have been identified in *Drosophila melanogaster* (D-IAP and D-IAP-2) and several in rodents and humans including human AAC-11 protein, X-linked inhibitor of apoptosis, XIAP and an XIAP associated protein ZAP-1 (Liston, P. et al. (1997) *Apoptosis* 2(5):423-441 and Liston, P. et al. (1996) *Nature* 379:349-352). Several DAD1 (defender against death) proteins have also been identified in mammalian and plant cells (Gallosis, P. et al. (1997) *Plant J.* 11(6):1325-1331). These proteins appear to be a subunit of oligosaccharyltransferase (OST), an enzyme that is involved in N-linked glycosylation in eukaryotes and may also be associated with programmed cell death in mammalian and plant cells. No genes encoding IAP, IAP-2, ZAP-1 or AAC-11 proteins in plants, especially corn, rice soybean and wheat, have been isolated and sequenced.

25

There is a great deal of interest in identifying the genes that encode IAPs in plants. These genes may be used to express IAPs in plant cells to enhance cell tissue culture growth. Accordingly, the availability of nucleic acid sequences encoding all or a portion of IAPs would facilitate studies to better understand programmed cell death in plants, provide genetic tools to enhance cell growth in tissue culture, increase the efficiency of gene transfer and help provide more stable transformations.

30

SUMMARY OF THE INVENTION

The instant invention relates to isolated nucleic acid fragments encoding proteins that inhibit apoptosis. Specifically, this invention concerns an isolated nucleic acid fragment encoding an XIAP associated factor-1 protein. In addition, this invention relates to a nucleic acid fragment that is complementary to the nucleic acid fragment encoding an XIAP associated factor-1 protein. An additional embodiment of the instant invention pertains to a polypeptide encoding all or a substantial portion of an XIAP associated factor-1 protein.

35

In another embodiment, the instant invention relates to a chimeric gene encoding an XIAP associated factor-1 protein, or to a chimeric gene that comprises a nucleic acid fragment that is complementary to a nucleic acid fragment encoding an XIAP associated factor-1 protein, operable linked to suitable regulatory sequences, wherein expression of the chimeric gene results in production of levels of the encoded protein in a transformed host cell that is altered (i.e., increased or decreased) from the level produced in an untransformed host cell.

In a further embodiment, the instant invention concerns a transformed host cell comprising in its genome a chimeric gene encoding an XIAP associated factor-1 protein, operably linked to suitable regulatory sequences. Expression of the chimeric gene results in production of altered levels of the encoded protein in the transformed host cell. The transformed host cell can be of eukaryotic or prokaryotic origin, and include cells derived from higher plants and microorganisms. The invention also includes transformed plants that arise from transformed host cells of higher plants, and seeds derived from such transformed plants.

An additional embodiment of the instant invention concerns a method of altering the level of expression of an XIAP associated factor-1 protein in a transformed host cell comprising: a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding an XIAP associated factor-1 protein; and b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of altered levels of XIAP associated factor-1 protein in the transformed host cell.

An addition embodiment of the instant invention concerns a method for obtaining a nucleic acid fragment encoding all or a substantial portion of an amino acid sequence encoding an XIAP associated factor-1 protein.

BRIEF DESCRIPTION OF THE FIGURES AND SEQUENCE LISTINGS

The invention can be more fully understood from the following detailed description and the accompanying Sequence Listing which form a part of this application.

Figure 1 shows a comparison of the amino acid sequences of the human XIAP associated factor-1 protein (SEQ ID NO:19) and SEQ ID NOs:4, 10, 12, 14 and 16.

The following sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

SEQ ID NO:1 is the nucleotide sequence comprising a contig assembled from the cDNA inserts in clones cbn2.pk0002.d12, chpc8.pk0002.d2 and p0093.cssap81r encoding a portion of a corn XIAP associated factor-1 protein.

SEQ ID NO:2 is the deduced amino acid sequence of a portion of an XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:1.

SEQ ID NO:3 is the nucleotide sequence comprising a contig assembled from the cDNA inserts in clones cr1n.pk0004.d9 and cs1l.pk0014.d5 encoding an entire corn XIAP associated factor-1 protein.

5 SEQ ID NO:4 is the deduced amino acid sequence of an entire XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:3.

SEQ ID NO:5 is the nucleotide sequence comprising a portion of the cDNA insert in clone ctal1n.pk0046.b2 encoding a portion of a corn XIAP associated factor-1 protein.

SEQ ID NO:6 is the deduced amino acid sequence of a portion of an XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:5.

10 SEQ ID NO:7 is the nucleotide sequence comprising a portion of the cDNA insert in clone rls72.pk0005.f3 encoding a portion of a rice XIAP associated factor-1 protein.

SEQ ID NO:8 is the deduced amino acid sequence of a portion of an XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:7.

15 SEQ ID NO:9 is the nucleotide sequence comprising a portion of the cDNA insert in clone rr1.pk0064.c1 encoding a entire rice XIAP associated factor-1 protein.

SEQ ID NO:10 is the deduced amino acid sequence of an entire XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:9.

SEQ ID NO:11 is the nucleotide sequence comprising the entire cDNA insert in clone ses2w.pk0023.e4 encoding an entire soybean XIAP associated factor-1 protein.

20 SEQ ID NO:12 is the deduced amino acid sequence of an entire XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:11.

SEQ ID NO:13 is the nucleotide sequence comprising a contig assembled from the cDNA inserts in clones sfl1.pk0123.b8 and sl1.pk147.e6 encoding a portion of an XIAP associated factor-1 protein.

25 SEQ ID NO:14 is the deduced amino acid sequence of a portion of an XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:13.

SEQ ID NO:15 is the nucleotide sequence comprising the entire cDNA insert in clone wdk3c.pk005.k15 encoding an entire wheat XIAP associated factor-1 protein.

30 SEQ ID NO:16 is the deduced amino acid sequence of an entire XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:15.

SEQ ID NO:17 is the nucleotide sequence comprising a portion of the cDNA insert in clone wle1n.pk0050.e8 encoding a portion of an XIAP associated factor-1 protein.

SEQ ID NO:18 is the deduced amino acid sequence of a portion of an XIAP associated factor-1 protein derived from the nucleotide sequence of SEQ ID NO:17.

35 SEQ ID NO:19 is the amino acid sequence of a human XIAP associated factor-1 sequence set forth in NCBI Identifier No. gi 1869901.

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Research* 13:3021-3030 (1985) and in the *Biochemical*

Journal 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF THE INVENTION

5 In the context of this disclosure, a number of terms shall be utilized. As used herein, an "isolated nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA. As used herein, "contig" refers
10 to an assemblage of overlapping nucleic acid sequences to form one contiguous nucleotide sequence. For example, several DNA sequences can be compared and aligned to identify common or overlapping regions. The individual sequences can then be assembled into a single contiguous nucleotide sequence.

As used herein, "substantially similar" refers to nucleic acid fragments wherein
15 changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the protein encoded by the DNA sequence.

"Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by antisense or co-suppression technology. "Substantially
20 similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate alteration of gene expression by antisense or co-suppression technology or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention
25 encompasses more than the specific exemplary sequences.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a gene which
30 result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded protein, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one
35 negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the protein molecule would also not be expected to alter the activity of the protein. Each of the proposed modifications is well within the

routine skill in the art, as is determination of retention of biological activity of the encoded products.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize, under stringent conditions (0.1X SSC, 0.1% SDS, 65°C), with the nucleic acid fragments disclosed herein.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent similarity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Preferred are those nucleic acid fragments whose nucleotide sequences encode amino acid sequences that are 80% similar to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% similar to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% similar to the amino acid sequences reported herein. Sequence alignments and percent similarity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins, D. G. and Sharp, P. M. (1989) *CABIOS* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to afford putative identification of that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches partial or complete amino acid and nucleotide sequences encoding one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may

now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

5 “Codon degeneracy” refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment that encodes all or a substantial portion of the amino acid sequence encoding the XIAP associated factor-1 proteins as set forth in SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16 and 18. The skilled artisan is well aware of the “codon-bias” exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

15 “Synthetic genes” can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form gene segments which are then enzymatically assembled to construct the entire gene. “Chemically synthesized”, as related to a sequence of DNA, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of DNA may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the genes can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

25 “Gene” refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. “Native gene” refers to a gene as found in nature with its own regulatory sequences. “Chimeric gene” refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature.

30 “Endogenous gene” refers to a native gene in its natural location in the genome of an organism. A “foreign” gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A “transgene” is a gene that has been introduced into the genome by a transformation procedure.

“Coding sequence” refers to a DNA sequence that codes for a specific amino acid sequence. “Regulatory sequences” refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

“Promoter” refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an “enhancer” is a DNA sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as “constitutive promoters”. New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg, (1989) *Biochemistry of Plants* 15:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

The “translation leader sequence” refers to a DNA sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner, R. and Foster, G. D. (1995) *Molecular Biotechnology* 3:225).

The “3' non-coding sequences” refer to DNA sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al., (1989) *Plant Cell* 1:671-680.

“RNA transcript” refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is

referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to RNA transcript that includes the mRNA and so can be translated into protein by the cell. "Antisense RNA" refers to a RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (U.S. Pat. No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Pat. No. 5,231,020, incorporated herein by reference).

"Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide. A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels, J. J., (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.*

42:21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal (*supra*) can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal (*supra*) may be added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal included

5 (Raikhel (1992) *Plant Phys.* 100:1627-1632).

“Transformation” refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as “transgenic” organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol.* 143:277) and particle-accelerated or “gene gun” transformation technology (Klein et al. (1987) *Nature (London)* 327:70-73; U.S. Pat. No. 4,945,050, incorporated herein by reference).

10

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter “Maniatis”).

15

Nucleic acid fragments encoding at least a portion of several proteins that inhibit apoptosis have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. Table 1 lists the proteins that are described herein, and the designation of the cDNA clones that comprise the nucleic acid fragments encoding these proteins.

20

TABLE 1
Apoptosis Inhibitory Proteins

25

Enzyme	Clone	Plant
XIAP associated factor-1	cbn2.pk0002.d12	Corn
	chpc8.pk0002.d2	Corn
	cr1n.pk0004.d9	Corn
	csi1.pk0014.d5	Corn
	cta1n.pk0046.b2	Corn
	p0093.cssap81r	Corn
	rls72.pk0005.f3	Rice
	rr1.pk0064.c1	Rice
	ses2w.pk0023.e4	Soybean
	sfl1.pk0123.b8	Soybean
	sl1.pk147.e6	Soybean
	wle1n.pk0050.e8	Wheat
	wdk3c.pk005.k15	Wheat

The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding other XIAP associated factor-1 proteins, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al., (1988) *PNAS USA* 85:8998) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al., (1989) *PNAS USA* 86:5673; Loh et al., (1989) *Science* 243:217). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman, M.A. and Martin, G.R., (1989) *Techniques* 1:165).

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing

portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest

5 (Lerner, R. A. (1984) *Adv. Immunol.* 36:1; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed XIAP associated factor-1 proteins are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of XIAP associated factor-1
10 proteins in those cells. Altering the level of XIAP associated factor-1 proteins in cells could facilitate studies to better understand programmed cell death in plants, provide genetic tools to alter programmed cell death, enhance cell growth in tissue culture, increase the efficiency of gene transfer and help provide more stable transformations.

Overexpression of the XIAP associated factor-1 proteins of the instant invention may
15 be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The
20 instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then be constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the
25 plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al., (1985) *EMBO J.* 4:2411-2418; De Almeida et al., (1989) *Mol. Gen. Genetics* 218:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired
30 expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

For some applications it may be useful to direct the instant apoptosis inhibitory proteins to different cellular compartments, or to facilitate its secretion from the cell. It is
35 thus envisioned that the chimeric gene described above may be further supplemented by altering the coding sequence to encode XIAP associated factor-1 proteins with appropriate intracellular targeting sequences such as transit sequences (Keegstra, K. (1989) *Cell* 56:247-253), signal sequences or sequences encoding endoplasmic reticulum localization (Chrispeels, J. J., (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53), or nuclear

localization signals (Raikhel, N. (1992) *Plant Phys.* 100:1627-1632) added and/or with targeting sequences that are already present removed. While the references cited give examples of each of these, the list is not exhaustive and more targeting signals of utility may be discovered in the future.

5 It may also be desirable to reduce or eliminate expression of genes encoding XIAP associated factor-1 proteins in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant apoptosis inhibitory proteins can be constructed by linking a gene or gene fragment encoding an XIAP associated factor-1 protein to plant promoter sequences. Alternatively, a chimeric gene designed to express
10 antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

15 The instant XIAP associated factor-1 proteins (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting XIAP associated factor-1 proteins *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant XIAP
20 associated factor-1 proteins are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct a chimeric gene for production of the instant XIAP associated factor-1 proteins. This chimeric gene could then be introduced into appropriate microorganisms via
25 transformation to provide high level expression of the encoded apoptosis inhibitory protein. An example of a vector for high level expression of the instant XIAP associated factor-1 proteins in a bacterial host is provided (Example 6).

30 All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then
35 be subjected to genetic analyses using computer programs such as MapMaker (Lander et al., (1987) *Genomics* 1:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted

and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein, D. et al., (1980) *Am. J. Hum. Genet.* 32:314-331).

5 The production and use of plant gene-derived probes for use in genetic mapping is described in R. Bernatzky, R. and Tanksley, S. D. (1986) *Plant Mol. Biol. Reporter* 4(1):37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to
10 those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; see Hoheisel, J. D., et al., In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

15 In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask, B. J. (1991) *Trends Genet.* 7:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan, M. et al. (1995) *Genome Research* 5:13-20), improvements in sensitivity may allow performance of FISH mapping
20 using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian, H. H. (1989) *J. Lab. Clin. Med.* 114(2):95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield, V. C. et al. (1993) *Genomics* 25 16:325-332), allele-specific ligation (Landegren, U. et al. (1988) *Science* 241:1077-1080), nucleotide extension reactions (Sokolov, B. P. (1990) *Nucleic Acid Res.* 18:3671), Radiation Hybrid Mapping (Walter, M. A. et al. (1997) *Nature Genetics* 7:22-28) and Happy Mapping (Dear, P. H. and Cook, P. R. (1989) *Nucleic Acid Res.* 17:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in
30 the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

35 Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer, (1989) *Proc. Natl. Acad. Sci USA* 86:9402; Koes et al., (1995) *Proc. Natl. Acad. Sci USA* 92:8149; Bensen et al., (1995) *Plant Cell* 7:75). The latter approach may be

accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The

5 amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the XIAP associated factor-1 protein. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as

10 that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding an XIAP associated factor-1 protein can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the XIAP associated factor-1 protein gene product.

EXAMPLES

15 The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without

20 departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Composition of cDNA Libraries; Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various corn, rice, soybean and wheat

25 tissues were prepared. The characteristics of the libraries are described below.

TABLE 2

cDNA Libraries from Corn, Rice, Soybean and Wheat

Library	Tissue	Clone
cbn2	Corn (<i>Zea mays</i> L.) developing kernel two days after pollination	cbn2.pk0002.d12
chpc8	Corn (<i>Zea mays</i> L., MBS847) 8 day old shoot treated with herbicide for 8 hrs**	chpc8.pk0002.d2
crln	Corn (<i>Zea mays</i> L.) root from 7 day seedlings grown in light*	cr1n.pk0004.d9
csi1	Corn (<i>Zea mays</i> L.) silk	csi1.pk0014.d5
cta1n	Corn (<i>Zea mays</i> L.) tassel*	cta1n.pk0046.b2
p0093	Corn stalk and shank at the intersection of these tissues (includes some internodal stalk and shank tissues) Growth conditions: field, 2-3 weeks after pollen shed; plants were allowed to pollinate naturally	p0093.cssap81r

rls72	Rice (<i>Oryza sativa</i> L.) leaf (15 days after germination) 72 hours after infection of <i>Magaporthe grisea</i> 4360-R-67 (avr2-yamo); Susceptible	rls72.pk0005.f3
rr1	Rice (<i>Oryza sativa</i> L.) root of two week old developing seedling	rr1.pk0064.c1
ses2w	Soybean (<i>Glycine max</i> L.) embryogenic suspension 2 weeks after subculture	ses2w.pk0023.e4
sfl1	Soybean (<i>Glycine max</i> L.) immature flower	sfl1.pk0123.b8
sl1	Soybean (<i>Glycine max</i> L.) Two week old developing seedlings treated with water	sl1.pk147.e6
wdk3c	Wheat (<i>Triticum aestivum</i> L.) developing kernel, 14 days after anthesis	wdk3c.pk005.k15
wle1n	Wheat (<i>Triticum aestivum</i> L.) leaf 7 day old etiolated seedling*	wle1n.pk0050.e8

* These libraries were normalized essentially as described in U.S. Pat. No. 5,482,845

** Application of 2-[(2,4-dihydro-2,6,9-trimethyl[1]benzothiopyrano[4,3-c]pyrazol-8-yl)carbonyl]-1,3-cyclohexanedione S,S-dioxide; synthesis and methods of using this compound are described in WO 97/19087, incorporated herein by reference.

5

cDNA libraries were prepared in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). Conversion of the Uni-ZAP™ XR libraries into plasmid libraries was accomplished according to the protocol provided by Stratagene. Upon conversion, cDNA inserts were contained in the plasmid vector pBluescript. cDNA inserts from randomly picked bacterial colonies containing recombinant pBluescript plasmids were amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences or plasmid DNA was prepared from cultured bacterial cells. Amplified insert DNAs or plasmid DNAs were sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams, M. D. et al., (1991) *Science* 252:1651). The resulting ESTs were analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

15

EXAMPLE 2

Identification of cDNA Clones

ESTs encoding apoptosis inhibitory proteins were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information

25

(NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish, W. and States, D. J. (1993) *Nature Genetics* 3:266-272 and Altschul, Stephen F., et al. (1997) *Nucleic Acids Res.* 25:3389-3402) provided by the NCBI.

For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3

Characterization of cDNA Clones Encoding XIAP Associated Factor-1 Homologs

The BLASTX search using the nucleotide sequences from clones cr1n.pk0004.d9, csil.pk0014.d5, cbn2.pk0002.d12, chpc8.pk0002.d2, ctal1n.pk0046.b2, p0093.cssap81r, rls72.pk0005.f3, rr1.pk0064.c1, sfl1.pk0123.b8, ses2w.pk0023.e4, sl1.pk147.e6, wdk3c.pk0005.k15 and wle1n.pk0050.e8 revealed similarity of the proteins encoded by the cDNAs to XIAP associated factor-1 from *Homo sapiens* (NCBI Identifier No. gi 1869901).

In the process of comparing the ESTs it was found that several clones had overlapping regions of homology. Using this homology it was possible to align the ESTs and assemble several individual contigs encoding unique plant XIAP associated factor-1 proteins. The EST composition of each contig is shown in Table 3. The BLAST results for each of the assembled contigs and individual ESTs and are also shown in Table 3:

TABLE 3

BLAST Results for Clones Encoding Polypeptides Homologous
to *Homo sapiens* XIAP Associated Factor-1 Proteins

Clone	BLAST pLog Score
Contig composed of: cbn2.pk0002.d12 chpc8.pk0002.d2 p0093.cssap81r	23.04
Contig composed of: cr1n.pk0004.d9 csil.pk0014.d5	28.30
ctal1n.pk0046.b2	3.70
rls72.pk0005.f3	13.52
rr1.pk0064.c1	26.30
ses2w.pk0023.e4	27.70

Clone	BLAST pLog Score
Contig composed of: sfl1.pk0123.b8 sl1.pk147.e6	29.70
wdk3c.pk005.k15	20.22
wle1n.pk0050.e8	10.40

The sequence of the corn contig composed of clones cbn2.pk0002.d12, chpc8.pk0002.d2 and p0093.cssap81r is shown in SEQ ID NO:1; the deduced amino acid sequence of this contig, which represents 60% of an XIAP associated factor-1 protein is shown in SEQ ID NO:2. The sequence of the corn contig composed of clones cr1n.pk0004.d9 and cs1l.pk0014.d5 is shown in SEQ ID NO:3; the deduced amino acid sequence of this contig, which represents 100% of the XIAP associated factor-1 protein is shown in SEQ ID NO:4. The amino acid sequence set forth in SEQ ID NO:4 was evaluated by BLASTP, yielding a pLog value of 24.52 versus the human XIAP sequence. The sequence of a portion of the cDNA insert from clone cta1n.pk0046.b2 is shown in SEQ ID NO:5; the deduced amino acid sequence of this cDNA, the N-terminal 65 amino acids of an XIAP associated factor-1, is shown in SEQ ID NO:6.

The sequence of a portion of the cDNA insert from clone rls72.pk0005.f3 is shown in SEQ ID NO:7; the deduced amino acid sequence of this cDNA, which represents the N-terminal 55 amino acids of the XIAP associated factor-1 protein, is shown in SEQ ID NO:8. The sequence of the entire cDNA insert in clone rr1.pk0064.c1 was determined and is shown in SEQ ID NO:9; the deduced amino acid sequence of this cDNA, which represents 100% of the XIAP associated factor-1 protein, is shown in SEQ ID NO:10. The amino acid sequence set forth in SEQ ID NO:10 was evaluated by BLASTP, yielding a pLog value of 23.15 versus the human sequence.

The sequence of the entire cDNA insert in clone ses2w.pk0023.e4 was determined and is shown in SEQ ID NO:11; the deduced amino acid sequence of this cDNA, which represents 100% of the XIAP associated factor-1 protein, is shown in SEQ ID NO:12. The amino acid sequence set forth in SEQ ID NO:12 was evaluated by BLASTP, yielding a pLog value of 24.52 versus the human sequence. The sequence of the soybean contig composed of clones sfl1.pk0123.b8 and sl1.pk147.e6 is shown in SEQ ID NO:13; the deduced amino acid sequence of this contig is shown in SEQ ID NO:14. The amino acid sequence of the contig represents the N-terminal 90% of an XIAP associated factor-1 protein. The amino acid sequence set forth in SEQ ID NO:14 was evaluated by BLASTP, yielding a pLog value of 24.70 versus the human XIAP associated factor-1 sequence.

The sequence of a portion of the cDNA insert from clone wdk3c.pk005.k15 is shown in SEQ ID NO:15; the deduced amino acid sequence of this cDNA is shown in SEQ ID

NO:16. The amino acid sequence of clone wdk3c.pk005.k15 represents 100% of the XIAP associated factor protein. The amino acid sequence set forth in SEQ ID NO:16 was evaluated by BLASTP, yielding a pLog value of 19.52 versus the human XIAP associated factor-1 sequence. The sequence of a portion of the cDNA insert from clone
 5 wle1n.pk0050.e8 is shown in SEQ ID NO:17; the deduced amino acid sequence of this cDNA, which represents 36% of the XIAP associated factor-1 protein, is shown in SEQ ID NO:18.

Figure 1 presents an alignment of the entire or nearly entire XIAP associated factor-1 amino acid sequences set forth in SEQ ID NOs:4, 10, 12, 14 and 16 and the human XIAP
 10 associated factor-1 sequence. The data in Table 4 represents a calculation of the percent similarity of the amino acid sequences set forth in SEQ ID NOs:4, 10, 12, 14 and 16 and the human XIAP associated factor-1 sequence.

TABLE 4

15 Percent Similarity of Amino Acid Sequences Deduced From the Nucleotide Sequences of cDNA Clones Encoding Polypeptides Homologous to *Homo sapiens* XIAP Associated Factor-1

Clone	SEQ ID NO.	Percent Similarity
Contig composed of: cr1n.pk0004.d9 cs1l.pk0014.d5	4	23.7%
rr1.pk0064.c1	10	23.4%
ses2w.pk0023.e4	12	26.0%
Contig composed of: sfl1.pk0123.b8 sl1.pk147.e6	14	28.2%
wdk3c.pk005.k15	16	22.5%

The percent similarity between each of the assembled contigs and the individual ESTs
 20 shown in Table 3, ranged between 30% and 77%.

Sequence alignments and percent similarity calculations were performed by the Clustal Algorithm (Higgins, D. G. et al., (1989) *CABIOS* 5(2):151-153), using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Default parameters for the Clustal method for protein multiple alignments are: GAP
 25 PENALTY=10, GAP LENGTH PENALTY=10; for pairwise alignments KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

Sequence alignments and BLAST scores and probabilities indicate that the instant nucleic acid fragments encode entire or nearly entire XIAP associated factor-1 proteins. These sequences represent the first plant sequences encoding XIAP associated factor-1
 30 proteins.

EXAMPLE 4

Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding apoptosis inhibitory protein in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below.

Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb Sall-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-Sall fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U. S. Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding an apoptosis inhibitory protein, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al., (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers

resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

5 The particle bombardment method (Klein et al., (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μ m in diameter) are coated with DNA using the following technique. Ten μ g of plasmid DNAs are added to 50 μ L of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μ L of a 2.5 M solution) and spermidine free base (20 μ L of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μ L of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μ L of ethanol. An aliquot (5 μ L) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

20 For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

25 Seven days after bombardment the tissue can be transferred to N6 medium that contains glufosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing glufosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

35 Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al., (1990) *Bio/Technology* 8:833-839).

EXAMPLE 5

Expression of Chimeric Genes in Dicot Cells

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem.* 261:9228-9238) can be used for expression of the instant apoptosis inhibitory proteins in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embryos may then be transformed with the expression vector comprising sequences encoding an apoptosis inhibitory protein. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can be maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Kline et al. (1987) *Nature* (London) 327:70, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al. (1983) *Gene* 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the apoptosis inhibitory

protein and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 μ L of a 60 mg/mL 1 μ m gold particle suspension is added (in order): 5 μ L DNA (1 μ g/ μ L), 20 μ L spermidine (0.1 M), and 50 μ L CaCl_2 (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 μ L 70% ethanol and resuspended in 40 μ L of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five μ L of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

EXAMPLE 6

Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant apoptosis inhibitory proteins can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene* 56:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve

GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 µg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 µL of water.

- 5 Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can
- 10 then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 µg/mL ampicillin. Transformants containing the gene encoding the apoptosis inhibitory protein are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

- 15 For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol.* 189:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio-β-galactoside, the inducer) can be added to a final concentration of
- 20 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 µL of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant
- 25 determined. One µg of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

CLAIMS

What is claimed is:

1. An isolated nucleic acid fragment encoding all or a substantial portion of an XIAP associated factor-1 protein comprising a member selected from the group consisting
5 of:
 - (a) an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16 and 18;
 - 10 (b) an isolated nucleic acid fragment that is substantially similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16 and 18; and
 - (c) an isolated nucleic acid fragment that is complementary to (a) or (b).
2. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid
15 fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:2.
3. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:4.
- 20 4. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:6.
5. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding
25 all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:8.
6. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:10.
7. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid
30 fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:12.
8. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:14.
- 35 9. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:16.

10. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment hybridizes under stringent conditions to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:18.
- 5 11. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:2.
12. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:4.
- 10 13. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:6.
14. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:8.
- 15 15. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:10.
16. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:12.
- 20 17. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:14.
- 25 18. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:16.
19. The isolated nucleic acid fragment of Claim 1 wherein the isolated nucleic acid fragment is at least 80% similar to an isolated nucleic acid fragment encoding all or a substantial portion of the amino acid sequence set forth in SEQ ID NO:18.
- 30 20. The isolated nucleic acid fragment of Claim 1 wherein the nucleotide sequence of the fragment comprises all or a portion of the sequence set forth in a member selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15 and 17.
21. A chimeric gene comprising the nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences.
- 35 22. A transformed host cell comprising the chimeric gene of Claim 21.
23. An XIAP associated factor-1 polypeptide comprising all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16 and 18.

24. A method of altering the level of expression of an apoptosis inhibitory protein in a host cell comprising:

- (a) transforming a host cell with the chimeric gene of Claim 21; and
- (b) growing the transformed host cell produced in step (a) under conditions that are suitable for expression of the chimeric gene

wherein expression of the chimeric gene results in production of altered levels of an apoptosis inhibitory protein in the transformed host cell.

25. A method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding an apoptosis inhibitory protein comprising:

- (a) probing a cDNA or genomic library with the nucleic acid fragment of Claim 1;
- (b) identifying a DNA clone that hybridizes with the nucleic acid fragment of Claim 1;
- (c) isolating the DNA clone identified in step (b); and
- (d) sequencing the cDNA or genomic fragment that comprises the clone isolated in step (c)

wherein the sequenced nucleic acid fragment encodes all or a substantial portion of the amino acid sequence encoding an apoptosis inhibitory protein.

26. A method of obtaining a nucleic acid fragment encoding a substantial portion of an amino acid sequence encoding an apoptosis inhibitory protein comprising:

- (a) synthesizing an oligonucleotide primer corresponding to a portion of the sequence set forth in any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15 and 17; and
- (b) amplifying a cDNA insert present in a cloning vector using the oligonucleotide primer of step (a) and a primer representing sequences of the cloning vector

wherein the amplified nucleic acid fragment encodes a substantial portion of an amino acid sequence encoding an apoptosis inhibitory protein.

27. The product of the method of Claim 25.

28. The product of the method of Claim 26.

SEQ ID NO:19	MEGD-----FVCRNCKRHVVVSANFTLHEAYCLRFLVL-CP-----
SEQ ID NO:4	MAAAADSDTA-----AAGAMSTCAHCQREIPFLNIDLHSVHCARNLQ-KCQ-----
SEQ ID NO:10	MAAAATAADAGSDPSAALAAAATSTCAHCQREIPSSNIDLHSAHCARNLQ-KCE-----
SEQ ID NO:12	MEAVSDQD-----TSICTHCDRAIPAANIDLHYAHCSRKLE-KCKV-----
SEQ ID NO:14	MAAVSDQD-----TSVCTHCDRAIPSANIDLHFVHCSRNL-RCQI-----
SEQ ID NO:16	MAVAADSADPA-----VATSTCAHCNREIPSPNITLHSAHCARNLHVKCE-----
SEQ ID NO:19	61
SEQ ID NO:4	-----ECEEVPKETMEEHCKLEHQVVGCTMCQQSMQKSSLEF-----HKANECQ
SEQ ID NO:10	-----HCGEMI PRKLMDEHYDENHAPINCSLCKETIERESWDL-----HKGEKCP
SEQ ID NO:12	-----HCGEMVARKLMDEHYNESHAPVNVCTLCKEIVTREIWDL-----QKSEQCP
SEQ ID NO:14	-----CGDMVPRKNAEDHYLRTHAPVSCSLCSETMERDILDI-----HKGENCEP
SEQ ID NO:16	-----CDDMIPKLLAEHHYLNTHAPVACSLCSETMERDILDI-----HREENCP
	-----HCGDMVPRKHMDRHYDDNHAPVNCPRCEQVHEHELWDL-----HKRLQCP
SEQ ID NO:19	121
SEQ ID NO:4	E---RPVECKFCKLDMQLSKLELHESYCGSRTTELCCGCGQFIMHRMLAQHRDVCRSEQAQ
SEQ ID NO:10	Q---RIVACEYCEFFELPAVDLHEHQDVCGNRTELCTCRKYIRLREWTGHEIQCHTNSN-
SEQ ID NO:12	Q---RIVACEYCEFFELPAVDLHEHQDVCGNRTEFCQCTCKKYVRLREWIGHEIQCHANAN-
SEQ ID NO:14	Q---RIVTCEFFELPAIDLAHQEVCGNRTELCHLCNKYVRLRERFSHEARC-----N-
SEQ ID NO:16	K---RMITCEFFELPAIDLAHQEVCGNRTELCHLCNKYVRLRELYNHEDSC-----N-
	Q---RMLACQYCKFFELPADDIFEHQDVCGNRRTKYCQPCNKDILRLREWIGHEL--LLHSK-
SEQ ID NO:19	181
SEQ ID NO:4	-----240
SEQ ID NO:10	-----
SEQ ID NO:12	-----
SEQ ID NO:14	-----
SEQ ID NO:16	-----

FIG. 1

2/4

SEQ ID NO:19	241	-----	300
SEQ ID NO:4		-----	
SEQ ID NO:10		-----	
SEQ ID NO:12		-----	
SEQ ID NO:14		-----	
SEQ ID NO:16		-----	
SEQ ID NO:19	301	-----	360
SEQ ID NO:4		-----	
SEQ ID NO:10		-----	
SEQ ID NO:12		-----	
SEQ ID NO:14		-----	
SEQ ID NO:16		-----	
SEQ ID NO:19	361	-----	420
SEQ ID NO:4		-----	
SEQ ID NO:10		-----	
SEQ ID NO:12		-----	
SEQ ID NO:14		-----	
SEQ ID NO:16		-----	
SEQ ID NO:19	421	-----	480
SEQ ID NO:4		-----	
SEQ ID NO:10		-----	
SEQ ID NO:12		-----	
SEQ ID NO:14		-----	
SEQ ID NO:16		-----	

-----LGKGERISAPERIY-----
-----GSADT-----SRAI-----
-----ANASQTSSARI-----
-----GIQDSSVGTSRNV-----
-----TIQDNSAGSSRYV-----
-----TNAAESSDRSM-----

-----CHYCNQMIPENK-----YFHHMGKCCPDSEFK-----KH

-----PVGNPEIL-----PSSLPSQAAENQSTMEKDV-----RP-----
-----PEREL-----RP-----
-----PEREV-----RP-----
-----REAE-----
-----RPAE-----
-----LEKE-----

FIG. 1 (Continued)

3/4

```

491      SEQ ID NO:19      540
      SEQ ID NO:4      ---KT---RS---INRFPLHSESSKKAPRSKNKTLDPL---
      SEQ ID NO:10     ---PP---PP---VRQ---
      SEQ ID NO:12     ---PP---P---
      SEQ ID NO:14     ---FE---QG---ARR---
      SEQ ID NO:16     ---RD---XG---ARR---
      ---EG---GRH---

541      SEQ ID NO:19      600
      SEQ ID NO:4      L---MSEPKP---RTSSPRGDKA---AYDILRRCS
      SEQ ID NO:10     ---A---RPARPAHASP---HKRLLETTIA
      SEQ ID NO:12     ---V---RPPRPMHGAQ---HKRLLETTIA
      SEQ ID NO:14     ---R---PPLPPQNDFS---TKRLLETTIA
      SEQ ID NO:16     ---X---PQNDFS---RKLFFTTIA
      ---KQVRPAFLK---HKQLLLTIA

601      SEQ ID NO:19      660
      SEQ ID NO:4      QCG---ILPLPILNQHEKCRWLASSKRKTSEKFQD
      SEQ ID NO:10     VTG---IAVMIGSI---LFQR---DEGF---
      SEQ ID NO:12     VTG---IAVMIGSI---LFQR---EESF---
      SEQ ID NO:14     ITG---IAVILGSF---FLQKKAESSDVH---
      SEQ ID NO:16     FTG---IAVLGSI---
      ITG---LAVLIGSI---LFQRRIGSSEAQ---
```

FIG. 1 (continued)

SEQ ID NO:19	661	LEKERY	720
SEQ ID NO:4		----	----
SEQ ID NO:10		----	----
SEQ ID NO:12		----	----
SEQ ID NO:14		----	----
SEQ ID NO:16		----	----
SEQ ID NO:19	721	RF	780
SEQ ID NO:4		----	----
SEQ ID NO:10		----	----
SEQ ID NO:12		----	----
SEQ ID NO:14		----	----
SEQ ID NO:16		----	----

FIG. 1 (Continued)

SEQUENCE LISTING

<110> E. I. DU PONT DE NEMOURS AND COMPANY
 <120> INHIBITORS OF APOPTOSIS PROTEINS IN PLANTS
 <130> BB-1148-A
 <150> 60/078,144
 <151> MARCH 16, 1998
 <160> 19
 <170> MICROSOFT WORD VERSION 7.0A

<210> 1
 <211> 882
 <212> DNA
 <213> Zea mays

<220>
 <221> unsure
 <222> (508)

<220>
 <221> unsure
 <222> (755)..(756)

<220>
 <221> unsure
 <222> (772)

<220>
 <221> unsure
 <222> (830)

<220>
 <221> unsure
 <222> (845)

<220>
 <221> unsure
 <222> (856)

<220>
 <221> unsure
 <222> (858)

<400> 1
 ctcttgtccg caccatcgcc tccgctccga ttccgagact cggcgcgggc gggcgagacg 60
 tggatctcga cggaggagac tccgaccatc gctccatgga ggccgcggac tctgaatccg 120
 tcatctccac cttcacctgc ccgactgtc agcgggaggt tccatcttca aacattgcat 180
 tgcacagcgc gcactgctgc cggaaacctc agaagtgcga acattgcgga gatatggttg 240
 ccagggaagct tatggatgaa cactacgatg aaaaccatgc tccgatgaat tgctcacttt 300
 gcaaacacac agtagaacgc gagctatggg atcttcatac aggcatacaa tgcccacaaa 360
 ggatgctcgc gtgccaatac tgtcagtttg aactgcctgc agttgatctt tttgaacatc 420
 aggatctgtg tggaaatcgg acagaatatt gtcaaagttg caggaaagtat atcaagactt 480
 ccgtgaatgg attgggacat gaaccantc catacgattt tggaatgctg catcaagaac 540
 tttcaaggta accttccaac tggatgctgc cagtgcattc attttcaaa acaacattca 600
 agctaaaaat cggcccgggg aagggaagag accgcctccc ggggtattata ttaagaagag 660
 acggaaaacat gggcccgcgc cagaaattcc cgaacctggg cccccacaa ctaacggccg 720
 actcaacgtc atctgcaacg ccaaccggag gtggnnacaa gaacctaccg anacgggtgg 780
 gaaacgaggg atttttaaca agctgaaatc gcccgaggga tcaaccagan tgagggcacc 840
 cgagnctaca caggtanacc ttcaaattac taagggtata at 882

<210> 2
 <211> 127
 <212> PRT
 <213> Zea mays

<400> 2
 Met Glu Ala Ala Asp Ser Glu Ser Val Ile Ser Thr Phe Thr Cys Pro
 1 5 10 15
 His Cys Gln Arg Glu Val Pro Ser Ser Asn Ile Ala Leu His Ser Ala
 20 25 30
 His Cys Ala Arg Asn Leu Gln Lys Cys Glu His Cys Gly Asp Met Val
 35 40 45
 Ala Arg Lys Leu Met Asp Glu His Tyr Asp Glu Asn His Ala Pro Met
 50 55 60
 Asn Cys Ser Leu Cys Lys His Thr Val Glu Arg Glu Leu Trp Asp Leu
 65 70 75 80
 His Thr Gly Ile Gln Cys Pro Gln Arg Met Leu Ala Cys Gln Tyr Cys
 85 90 95
 Gln Phe Glu Leu Pro Ala Val Asp Leu Phe Glu His Gln Asp Leu Cys
 100 105 110
 Gly Asn Arg Thr Glu Tyr Cys Gln Ser Cys Arg Lys Tyr Ile Lys
 115 120 125

<210> 3
 <211> 1219
 <212> DNA
 <213> Zea mays

<400> 3
 gcacgagctc tcgctgccc a cccgatcaacc accccgcttc ccccgaccgc gtctcggagc 60
 acccttctca accgcatacct gcctctgctc ggtgtcgccg tctcgactgc aggagagggg 120
 agcgcgccgt cgccgatctg atcgggagcc ggaggccaag cgctcgaggtg agcccgtcag 180
 gaaccgccat ggcagcggcc gccgccgatt ccgacaccgc cgcagctgtg ggcgccatgt 240
 ccacctgctc ccactgtcag cgagaaatcc catttttgaa catcgacttg cactctgtac 300
 attgcgcccg caacctccaa aagtgccagc actgcggcga aatgattcct aggaaactta 360
 tggatgaaca ttatgatgaa aacctgccc cgatcaattg ctccctttgt aaagaaacaa 420
 tagaacgtga gagctgggat cttcataaag gtgaaaagtg cccacaaaga attgtcgctt 480
 gtgaatatgt cgagtttgaa ttgcctgcag ttgatctcca tgaacatcag gatgtatgtg 540
 gaaaccgaac agagctttgc caaacatgca ggaagtatat tagactgcgt gaatggacag 600
 ggcacgaaat acagtgccat acaaaactcaa atggttctgc agacacatcc cgagcaattc 660
 cagagagaga actacgaccg ccgccgccgc cagtgcgaca ggcgcgacca gcacgtcctg 720
 cccacgcttc accacacaag cggctcctct tcacaatcgc cgctcactgga atagcagtta 780
 tgatcggtc cactactgtt cagagggacg agggtttcta gctccagtag ccatggactc 840
 gctctgtcaa ggtgcttctt gaccgggaga gtcatgcggg catgccacag gaaggtggta 900
 ttaatgtata ataactcgtt cgaatcaatt acaataactg gagtttgta gggagtatag 960
 cccagggcat gcattcgtca aaaggcgaag ttatgttctg cccctggcag ttctgagact 1020
 tgtgtttgct ttgacactgc acttgaattt gttcatctgt aagtcagaga tgagttatac 1080
 aatcgccag tgtcatatct ggaattgcga cacgggtgcag cagctgttaa tactttggct 1140
 tatatcctgt aatgatactt tgctgttggg atgagtatag cgtatcaatt tgtggggaaa 1200
 aaaaaaaaaa aaaaatcca 1219

<210> 4
 <211> 210
 <212> PRT
 <213> Zea mays

<400> 4
 Met Ala Ala Ala Ala Asp Ser Asp Thr Ala Ala Ala Val Gly Ala
 1 5 10 15

Met Ser Thr Cys Ala His Cys Gln Arg Glu Ile Pro Phe Leu Asn Ile
 20 25 30
 Asp Leu His S r Val His Cys Ala Arg Asn Leu Gln Lys Cys Gln His
 35 40 45
 Cys Gly Glu Met Ile Pro Arg Lys Leu Met Asp Glu His Tyr Asp Glu
 50 55 60
 Asn His Ala Pro Ile Asn Cys Ser Leu Cys Lys Glu Thr Ile Glu Arg
 65 70 75 80
 Glu Ser Trp Asp Leu His Lys Gly Glu Lys Cys Pro Gln Arg Ile Val
 85 90 95
 Ala Cys Glu Tyr Cys Glu Phe Glu Leu Pro Ala Val Asp Leu His Glu
 100 105 110
 His Gln Asp Val Cys Gly Asn Arg Thr Glu Leu Cys Gln Thr Cys Arg
 115 120 125
 Lys Tyr Ile Arg Leu Arg Glu Trp Thr Gly His Glu Ile Gln Cys His
 130 135 140
 Thr Asn Ser Asn Gly Ser Ala Asp Thr Ser Arg Ala Ile Pro Glu Arg
 145 150 155 160
 Glu Leu Arg Pro Pro Pro Pro Pro Val Arg Gln Ala Arg Pro Ala Arg
 165 170 175
 Pro Ala His Ala Ser Pro His Lys Arg Leu Leu Phe Thr Ile Ala Val
 180 185 190
 Thr Gly Ile Ala Val Met Ile Gly Ser Ile Leu Phe Gln Arg Asp Glu
 195 200 205
 Gly Phe
 210

<210> 5
 <211> 599
 <212> DNA
 <213> Zea mays

<220>
 <221> unsure
 <222> (346)

<220>
 <221> unsure
 <222> (422)

<220>
 <221> unsure
 <222> (428)

<220>
 <221> unsure
 <222> (442)

<220>
 <221> unsur
 <222> (531)

<220>
 <221> unsure
 <222> (547)

<220>
 <221> unsure
 <222> (562)..(563)

<220>
 <221> unsure
 <222> (573)

<220>
 <221> unsure
 <222> (579)

<220>
 <221> unsure
 <222> (583)..(584)

<400> 5
 gtcctctgct ccgctccgca ccctcgccgc cccgaaattg ctcgactcgg cacaggcgag 60
 acgtggatct ccgcgaggga cctccgatca ccgatccatg gcggccgcgg actccgaccc 120
 cgccatctcc atctccacct gcccgactg tcagcgggag attccatctt caaacattgc 180
 attgcacagc gcgcactacg ctaggaacct tcaaaagtgc gaacattgcg gaaatatagt 240
 tccgagggaag cttgtggatc aacactacga tgaaaaccat gctccggcat cttgtttcc 300
 aaactccaag cttctcttag atgctgtcaa ttcataattg tagtantgtt caagtcaagc 360
 gttacattga gatgttggtt gatttacata tttgtatgt gatgggtaat tgtaacatgt 420
 anattgntt gtatacatat angaatgtac tacatattca aacggacaag ggggccggta 480
 ccaattcgcc tataggagtc tataccgcgc tcacggctcg tttaaactct nacgggaaac 540
 tgggtanac tatcgctgag anncccctcg cantggtana canngcccac gtccctcag 599

<210> 6
 <211> 65
 <212> PRT
 <213> Zea mays

<400> 6
 Met Ala Ala Ala Asp Ser Asp Pro Ala Ile Ser Ile Ser Thr Cys Pro
 1 5 10 15
 His Cys Gln Arg Glu Ile Pro Ser Ser Asn Ile Ala Leu His Ser Ala
 20 25 30
 His Tyr Ala Arg Asn Leu Gln Lys Cys Glu His Cys Gly Asn Ile Val
 35 40 45
 Pro Arg Lys Leu Val Asp Gln His Tyr Asp Glu Asn His Ala Pro Ala
 50 55 60

Ser
 65

<210> 7
 <211> 460
 <212> DNA
 <213> Oryza sativa

<220>
 <221> unsure
 <222> (308)

<220>
 <221> unsure
 <222> (438)

<220>
 <221> unsure
 <222> (452)

<220>
 <221> unsure
 <222> (457)

<400> 7
 tacagcctcc actccaatcc ctctccggac tcggtctctc tccccctacc tttttcagtt 60
 tttccccacc aaactcgcct cgccctcctc tcttccttct tcttcttcct cctccgagta 120
 ccagagcgca cgaaacccct cttegaattct ccgctctcct cccgcgggtg accccatccg 180
 accgatcgca agccatggcc gccgccgccg ccgactccgc cgatcccgtc atcaccaccg 240
 ccacctgcgc gcaactgccat cgagaaattc catctccaaa catcgcgctg catttgacac 300
 tgcgctcnaa ccttcaaaaag tgtgaacatt gtgggtatat gggtcaaaga agtaatggat 360
 gaacactatg atgaaacatg tcaatgattt gtcactttgc aaaaaatgta caacgtgagt 420
 atgggtcttc ataaaggnta catgccacaa gntgttnatg 460

<210> 8
 <211> 55
 <212> PRT
 <213> Oryza sativa

<400> 8
 Met Ala Ala Ala Ala Asp Ser Ala Asp Pro Val Ile Thr Thr Ala
 1 5 10 15

Thr Cys Ala His Cys His Arg Glu Ile Pro Ser Pro Asn Ile Ala Leu
 20 25 30

His Leu Thr Leu Arg Ser Asn Leu Gln Lys Cys Glu His Cys Gly Tyr
 35 40 45

Met Val Gln Arg Ser Asn Gly
 50 55

<210> 9
 <211> 551
 <212> DNA
 <213> Oryza sativa

<220>
 <221> unsure
 <222> (423)

<220>
 <221> unsure
 <222> (427)

<220>
 <221> unsure
 <222> (430)

<220>
 <221> unsure
 <222> (441)

<220>
 <221> unsure
 <222> (485)

<220>
 <221> unsure
 <222> (497)

<220>
 <221> unsur
 <222> (516) .. (517)

<220>
 <221> unsure
 <222> (520)

<220>
 <221> unsure
 <222> (526) .. (527)

<220>
 <221> unsure
 <222> (543)

<220>
 <221> unsure
 <222> (545)

<220>
 <221> unsure
 <222> (549)

<400> 9
 aactcctccc cccccccccc cccaaccaat cttcctcgaa tctcctcggt gcgccaccac 60
 ccaccccgcg cgcgctcgcg tcgtcttcgt cgtctccgcc gccggctgct cgattggcgc 120
 cggccaagga agcggcgagg cggattcact gcgaaggcga tatggcggca gcagcaacgg 180
 cggccgacgc cggtctccgat cctcctcgcg cctcgcgcgc cgccgcccgc acatccacct 240
 gcgcccactg ccagcgagaa attccatcgt caaacattga cttacattcg gcacactgcg 300
 cacgtaacct tcaaaagtgt gaacactgtg gagaaatggt ggccagaaaag cttatggacg 360
 aagcactaca atgaaagcca tgctccagta aattgcactc ttgcaagga aatttgtaac 420
 acnccganatn ggggattctc nataaaaggt gaacagtgcc cacaaaggga ttgttgcatg 480
 ttaanatgtg agttganttc ctgcggtgaa cttaannaan tcagannntg cggaacgaa 540
 cananttgna a 551

<210> 10
 <211> 217
 <212> PRT
 <213> *Oryza sativa*

<400> 10
 Met Ala Ala Ala Ala Thr Ala Ala Asp Ala Gly Ser Asp Pro Ser Ala
 1 5 10 15
 Ala Leu Ala Ala Ala Ala Thr Ser Thr Cys Ala His Cys Gln Arg
 20 25 30
 Glu Ile Pro Ser Ser Asn Ile Asp Leu His Ser Ala His Cys Ala Arg
 35 40 45
 Asn Leu Gln Lys Cys Glu His Cys Gly Glu Met Val Ala Arg Lys Leu
 50 55 60
 Met Asp Glu His Tyr Asn Glu Ser His Ala Pro Val Asn Cys Thr Leu
 65 70 75 80
 Cys Lys Glu Ile Val Thr Arg Glu Ile Trp Asp Leu Gln Lys Ser Glu
 85 90 95
 Gln Cys Pro Gln Arg Ile Val Ala Cys Glu Tyr Cys Glu Phe Glu Leu
 100 105 110
 Pro Ala Val Glu Leu His Glu His Gln Asp Val Cys Gly Asn Arg Thr
 115 120 125

Glu Phe Cys Gln Thr Cys Lys Lys Tyr Val Arg Leu Arg Glu Trp Ile
 130 135 140
 Gly His Glu Ile Gln Cys His Ala Asn Ala Asn Ala Asn Ala Ser Ala
 145 150 155 160
 Gln Thr Ser Ser Ala Arg Ile Ile Pro Glu Arg Glu Val Arg Pro Pro
 165 170 175
 Pro Pro Val Arg Pro Pro Arg Pro Met His Gly Ala Gln His Lys Arg
 180 185 190
 Leu Leu Phe Thr Ile Ala Val Thr Gly Ile Ala Val Met Ile Gly Ser
 195 200 205
 Ile Leu Phe Gln Arg Glu Glu Ser Phe
 210 215

<210> 11
 <211> 975
 <212> DNA
 <213> Glycine max

<400> 11
 tttttttttt ttttcctaaa actttcaatt attttcatca aatgcagagc tacagtcatt 60
 ttcaatgaag tatataacta aatatatatt ttataaaatg acaatgccga ttgtgttcta 120
 cacttctact acatcacatt agcaaaagta tatctcatcc ggcttctgat gcaggggaat 180
 ggattataca gaaaacgcag gatcctaaaa tgattcggct ttgttgctaa tgcacgtcac 240
 tggactctgc cttcttctgt aggaaaaaag atcccagtat gacagcaata ccagtgatgg 300
 ctattgtgaa aagaaggcgt ttagttgaga agtcgttctg cggcggcagt ggaggccttc 360
 ttctagcacc ttgctctctc tcagcttctc tcacattcct tgaagtgcc acagaactat 420
 cttggatgcc attgcatcga gcttcatggc tgaatcttcc acgaagtcta acatatttgt 480
 tacaaagggt acaaagtcca gttcgattcc cacatacttc ctgatgtca gccagatcaa 540
 tggctggcaa tggaactca cagaactcac aagtgacaat cctttgaggy caattttcac 600
 ctttatggat atctaaaaata tcacgctcca ttgtctcact gcataatgaa caagaaaccg 660
 gcgcatgggt tcttaaatag tgatcctccg cattttttct gggaaccatg tcaccacaaa 720
 ctttgcatth ttcaagcttc cgagagcaat gagcataatg caagtcaata tttgcagcag 780
 ggatagcccg atcacagtgg gtgcataac tagtgtcttg atccgaaaca gcttccatcg 840
 ccgaatatac acacaatcct acaccaaagg atccctctac gacaaaatct ccggaaacca 900
 aacaaaaaaa atagacggag aatttcggat tatacgagga ttattacggg gattaggttt 960
 tggacaacct cgtgc 975

<210> 12
 <211> 203
 <212> PRT
 <213> Glycine max

<400> 12
 Met Glu Ala Val Ser Asp Gln Asp Thr Ser Ile Cys Thr His Cys Asp
 1 5 10 15
 Arg Ala Ile Pro Ala Ala Asn Ile Asp Leu His Tyr Ala His Cys Ser
 20 25 30
 Arg Lys Leu Glu Lys Cys Lys Val Cys Gly Asp Met Val Pro Arg Lys
 35 40 45
 Asn Ala Glu Asp His Tyr Leu Arg Thr His Ala Pro Val Ser Cys Ser
 50 55 60
 Leu Cys Ser Glu Thr Met Glu Arg Asp Ile Leu Asp Ile His Lys Gly
 65 70 75 80
 Glu Asn Cys Pro Gln Arg Ile Val Thr Cys Glu Ph Cys Glu Phe Pro
 85 90 95

Leu Pro Ala Ile Asp Leu Ala Glu His Gln Glu Val Cys Gly Asn Arg
 100 105 110
 Thr Glu Leu Cys His Leu Cys Asn Lys Tyr Val Arg Leu Arg Glu Arg
 115 120 125
 Phe Ser His Glu Ala Arg Cys Asn Gly Ile Gln Asp Ser Ser Val Gly
 130 135 140
 Thr Ser Arg Asn Val Arg Glu Ala Glu Arg Glu Gln Gly Ala Arg Arg
 145 150 155 160
 Arg Pro Pro Leu Pro Pro Gln Asn Asp Phe Ser Thr Lys Arg Leu Leu
 165 170 175
 Phe Thr Ile Ala Ile Thr Gly Ile Ala Val Ile Leu Gly Ser Phe Phe
 180 185 190
 Leu Gln Lys Lys Ala Glu Ser Ser Asp Val His
 195 200

<210> 13
 <211> 795
 <212> DNA
 <213> Glycine max

<220>
 <221> unsure
 <222> (607)

<220>
 <221> unsure
 <222> (621)

<220>
 <221> unsure
 <222> (776)

<220>
 <221> unsure
 <222> (786)

<400> 13
 ttgtggttcaa caagaggaat tctcccaaatt tttgtatttt tcacccaaat ttttccacaa 60
 aacctaacc tctcttgtt tcacgcgcaa tccgaatctc catttctatc ttgtcctagg 120
 acgagcttgt taatacgcca tggcagccgt gtccgatcaa gacactagt tatgcactca 180
 ctgtgatcgg gctattcctt cggcaaatat tgatttgcatt tttgttcatt gctctcgaaa 240
 ccttgaaaga tgccaaattt gtgacgacat gattccgaaa ctacttgctg aggagcacta 300
 tttaaacaca catgccccgg ttgctgttc gctgtgcagt gaaaccatgg aacgtgatat 360
 tttagatatc catagagagg aaaattgtcc taaaaggatg atcacctgtg agttctgtga 420
 gtttcctttg cgggtattg atctggctga gcatcaggaa gtatgtggga atcgaacaga 480
 actttgtcac ctttgaaca aatatgttag actgcgtgaa ctatacaacc atgaagatag 540
 ttgcaataca attcaagaca attctgcagg gtcttcaagg tatgtgaggc cagctgaaag 600
 agatganggt gctagaagaa ngccgcagaa tgatttctca agaaagcgtc tttttttcac 660
 aatagcattc actggcattg ctgttgcct cggatcaatt ttttccagaa ggaagacaga 720
 tttcaacaat gttgcaatag caacgggaga agtatacgtg ttagtgtaaa tttgcntgat 780
 cgttanaatt tattt 795

<210> 14
 <211> 188
 <212> PRT
 <213> Glycine max

<220>
 <221> UNSURE
 <222> (156)

<220>
 <221> UNSURE
 <222> (161)

<400> 14
 Met Ala Ala Val Ser Asp Gln Asp Thr Ser Val Cys Thr His Cys Asp
 1 5 10 15
 Arg Ala Ile Pro Ser Ala Asn Ile Asp Leu His Phe Val His Cys Ser
 20 25 30
 Arg Asn Leu Glu Arg Cys Gln Ile Cys Asp Asp Met Ile Pro Lys Leu
 35 40 45
 Leu Ala Glu Glu His Tyr Leu Asn Thr His Ala Pro Val Ala Cys Ser
 50 55 60
 Leu Cys Ser Glu Thr Met Glu Arg Asp Ile Leu Asp Ile His Arg Glu
 65 70 75 80
 Glu Asn Cys Pro Lys Arg Met Ile Thr Cys Glu Phe Cys Glu Phe Pro
 85 90 95
 Leu Pro Ala Ile Asp Leu Ala Glu His Gln Glu Val Cys Gly Asn Arg
 100 105 110
 Thr Glu Leu Cys His Leu Cys Asn Lys Tyr Val Arg Leu Arg Glu Leu
 115 120 125
 Tyr Asn His Glu Asp Ser Cys Asn Thr Ile Gln Asp Asn Ser Ala Gly
 130 135 140
 Ser Ser Arg Tyr Val Arg Pro Ala Glu Arg Asp Xaa Gly Ala Arg Arg
 145 150 155 160
 Xaa Pro Gln Asn Asp Phe Ser Arg Lys Arg Leu Phe Phe Thr Ile Ala
 165 170 175
 Phe Thr Gly Ile Ala Val Val Leu Gly Ser Ile Phe
 180 185

<210> 15
 <211> 1024
 <212> DNA
 <213> Triticum aestivum

<400> 15
 gcacgagcac ctcaccacga gccccctcga taactcctcc aatccctctg ctcttctcgc 60
 ctccatcccc aacgaggacc ccctgattcg gcgagggtcc aattcggccg gcgactcggga 120
 tctcggcgcg ggggttgacg ggtgatccct caggactcgc aagccatggc cgtcgccgcc 180
 gactccgccc atcccgcctg cgccacctcc acctgcgcgc actgtaaccg agagattccg 240
 tctccgaaca tcacgctgca ttctgcacac tgtgctcgca accttcattg gaaatgtgag 300
 cattgtggag atatggtccc aaggaagcac atggataggc actacgacga taaccatgct 360
 ccggtgaatt gccacgctg cgaacaaact gtggagcatg agctgtggga tcttcataaa 420
 cgcttacagt gccacaaag gatgcttgcg tgccaatatt gcaagtttga actaccgcgc 480
 gatgacattt ttgagcatca ggatgtatgt ggaaatcgaa caaatattg tcaaccttgc 540
 aacaaggaca tcagacttgc cgaatggatt ggacacgagc tcctactcca ttcaaaaaca 600
 aatgctgctg cagaatcctc aagtgcacga tccatgctgg aaaaagaaga ggggtggtcgt 660
 cacaagcaag tacgcccggc ttttggttg aaacacaagc aactgctcct cacaattgca 720
 ataaccggac ttgctgtact gattggatcc attctgttcc aaagaaggat tggagttcc 780
 gaagcccagt gaagttgtaa tgatgttcac cgcaagaatg tagatcagtt tttccataat 840
 cttaatgcat aagttcttcc ataatcagta ggagagctta ttccttgagt aatgttcaga 900
 acacaacca tgtagaatat cagaaatcat ctcatcgatg tactgtgcct cacaatgtat 960
 gccaatatga tactccctcc gtagcttgca nnaaaaaa aataacaaaa aaaaaaaaaa 1020
 ataa 1024

<210> 16
 <211> 208
 <212> PRT
 <213> Triticum aestivum

<400> 16
 Met Ala Val Ala Ala Asp Ser Ala Asp Pro Ala Val Ala Thr Ser Thr
 1 5 10 15
 Cys Ala His Cys Asn Arg Glu Ile Pro Ser Pro Asn Ile Thr Leu His
 20 25 30
 Ser Ala His Cys Ala Arg Asn Leu His Val Lys Cys Glu His Cys Gly
 35 40 45
 Asp Met Val Pro Arg Lys His Met Asp Arg His Tyr Asp Asp Asn His
 50 55 60
 Ala Pro Val Asn Cys Pro Arg Cys Glu Gln Thr Val Glu His Glu Leu
 65 70 75 80
 Trp Asp Leu His Lys Arg Leu Gln Cys Pro Gln Arg Met Leu Ala Cys
 85 90 95
 Gln Tyr Cys Lys Phe Glu Leu Pro Ala Asp Asp Ile Phe Glu His Gln
 100 105 110
 Asp Val Cys Gly Asn Arg Thr Lys Tyr Cys Gln Pro Cys Asn Lys Asp
 115 120 125
 Ile Arg Leu Arg Glu Trp Ile Gly His Glu Leu Leu Leu His Ser Lys
 130 135 140
 Thr Asn Ala Ala Ala Glu Ser Ser Ser Asp Arg Ser Met Leu Glu Lys
 145 150 155 160
 Glu Glu Gly Gly Arg His Lys Gln Val Arg Pro Ala Phe Gly Leu Lys
 165 170 175
 His Lys Gln Leu Leu Leu Thr Ile Ala Ile Thr Gly Leu Ala Val Leu
 180 185 190
 Ile Gly Ser Ile Leu Phe Gln Arg Arg Ile Gly Ser Ser Glu Ala Gln
 195 200 205

<210> 17
 <211> 606
 <212> DNA
 <213> Triticum aestivum

<220>
 <221> unsure
 <222> (484)

<220>
 <221> unsure
 <222> (494)

<220>
 <221> unsure
 <222> (544)

<220>
 <221> unsure
 <222> (572)

<220>
 <221> unsure
 <222> (578)

<220>
 <221> unsure
 <222> (590)

<220>
 <221> unsure
 <222> (593)

<400> 17
 ccagcttttc cctcccccca ccaaacttcc accccttctt cgagacctcc ggccacccac 60
 ccgccaaggc ctccctccgc tccgcgctgt ccccgctgga ccaccgcagc actcccctgt 120
 ccgatcgga acgaggagag tgtgccacag gagagcgga tccgtccgtc cggccccaag 180
 acgagcgcca tggcagacgc cggctctcat atcgccgcgc tgcgcgccac ctccatctgc 240
 gccactgtc aacgagaaat tccatcgtca aacatcgact ttcatccgt aactgtgct 300
 cgtaaccttc aaaagtgcga agcactgtgg agaaatggtt cccaaggaag cttatggaaa 360
 gaacactaca atgaaaatca cgctccggtg aattgctccc tgtgcaaaga aaccttgcca 420
 acgggagata cttgatcttc acaaaagtga acatggcaca cagaggatgg ttgcgtgtgc 480
 atantgtgag tacnaattgc ctgcaattga tatcatgaac atcaaggatg ttatgtggaa 540
 acgnaacgga gtttgtcaaa catgcaagaa gntattanac tctgaatgn tanggcatga 600
 aatcat 606

<210> 18
 <211> 75
 <212> PRT
 <213> Triticum aestivum

<400> 18
 Met Ala Asp Ala Gly Ser His Ile Ala Ala Val Val Ala Thr Ser Ile
 1 5 10 15
 Cys Ala His Cys Gln Arg Glu Ile Pro Ser Ser Asn Ile Asp Phe His
 20 25 30
 Ser Val His Cys Ala Arg Asn Leu Gln Lys Cys Glu Ala Leu Trp Arg
 35 40 45
 Asn Gly Ser Gln Gly Ser Leu Trp Lys Glu His Tyr Asn Glu Asn His
 50 55 60
 Ala Pro Val Asn Cys Ser Leu Cys Lys Glu Thr
 65 70 75

<210> 19
 <211> 317
 <212> PRT
 <213> Homo sapiens

<400> 19
 Met Glu Gly Asp Phe Ser Val Cys Arg Asn Cys Lys Arg His Val Val
 1 5 10 15
 Ser Ala Asn Phe Thr Leu His Glu Ala Tyr Cys Leu Arg Phe Leu Val
 20 25 30
 Leu Cys Pro Glu Cys Glu Glu Pro Val Pro Lys Glu Thr Met Glu Glu
 35 40 45
 His Cys Lys Leu Glu His Gln Gln Val Gly Cys Thr Met Cys Gln Gln
 50 55 60

Ser Met Gln Lys Ser Ser Leu Glu Phe His Lys Ala Asn Glu Cys Gln
 65 70 75 80
 Glu Arg Pro Val Glu Cys Lys Phe Cys Lys Leu Asp Met Gln Leu Ser
 85 90 95
 Lys Leu Glu Leu His Glu Ser Tyr Cys Gly Ser Arg Thr Glu Leu Cys
 100 105 110
 Gln Gly Cys Gly Gln Phe Ile Met His Arg Met Leu Ala Gln His Arg
 115 120 125
 Asp Val Cys Arg Ser Glu Gln Ala Gln Leu Gly Lys Gly Glu Arg Ile
 130 135 140
 Ser Ala Pro Glu Arg Glu Ile Tyr Cys His Tyr Cys Asn Gln Met Ile
 145 150 155 160
 Pro Glu Asn Lys Tyr Phe His His Met Gly Lys Cys Cys Pro Asp Ser
 165 170 175
 Glu Phe Lys Lys His Phe Pro Val Gly Asn Pro Glu Ile Leu Pro Ser
 180 185 190
 Ser Leu Pro Ser Gln Ala Ala Glu Asn Gln Thr Ser Thr Met Glu Lys
 195 200 205
 Asp Val Arg Pro Lys Thr Arg Ser Ile Asn Arg Phe Pro Leu His Ser
 210 215 220
 Glu Ser Ser Ser Lys Lys Ala Pro Arg Ser Lys Asn Lys Thr Leu Asp
 225 230 235 240
 Pro Leu Leu Met Ser Glu Pro Lys Pro Arg Thr Ser Ser Pro Arg Gly
 245 250 255
 Asp Lys Ala Ala Tyr Asp Ile Leu Arg Arg Cys Ser Gln Cys Gly Ile
 260 265 270
 Leu Leu Pro Leu Pro Ile Leu Asn Gln His Gln Glu Lys Cys Arg Trp
 275 280 285
 Leu Ala Ser Ser Lys Arg Lys Thr Ser Glu Lys Phe Gln Leu Asp Leu
 290 295 300
 Glu Lys Glu Arg Tyr Tyr Lys Phe Lys Arg Phe His Phe
 305 310 315

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/05183

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/82 C07K14/415 C12N15/29

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	VYSOTSKAIA, V.S., ET AL.: "the sequence of BAC F21M12 from Arabidopsis thaliana chromosome 1" EMBL SEQUENCE DATA LIBRARY, 19 March 1997, XP002107428 heidelberg, germany accession no.AC000132	1-4,6-10
X	CARSON, D.L. AND BOTHA, F.C.: "sugarcane cDNA from leaf roll tissue" EMBL SEQUENCE DATA LIBRARY, 16 October 1996, XP002107429 heidelberg, germany accession no. AA 080664	1,3
	--- -/-	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

28 June 1999

Date of mailing of the international search report

12/07/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Holtorf, S